MISB MOTION IMAGERY STANDARDS BOARD	MISB TRM 0803.1
TECHNICAL REFERENCE MATERIAL **Delivery of Low Bandwidth Motion Imagery**	**27 February 2014**

1 Scope

This Motion Imagery Standards Board (MISB) TRM provides information for creating and distributing Motion Imagery (MI) over low bandwidth channels.

2 References

[1] MISP 6.6, Motion Imagery Standards Profile, Feb 2014
[2] MISB ST 1402 MPEG-2 Transport of Compressed Motion Imagery, Feb 2014
[3] ISO/IEC 13818-12013, Information Technology – Generic coding of moving pictures and associated audio information: Systems
[4] RFC 3550 RTP: A Transport Protocol for Real-Time Applications, Jul 2003
[5] MISB ST 0804.4 RTP for Motion Imagery and Metadata, Feb 2014
[6] RFC 2326, Real Time Streaming Protocol (RTSP), Apr 1998
[7] RFC 2327, SDP: Session Description Protocol, April 1998
[8] RFC 2616, Hypertext Transfer Protocol -- HTTP/1.1, 1999
[9] SMPTE ST 336:2007, Data Encoding Protocol Using Key-Length-Value
[10] IETF RFC 6597, RTP Payload Format for SMPTE ST 336 Encoded Data, Apr 2012
[11] MISB ST 1001.1 Audio Encoding, Feb 2014

3 Revision History

Revision	Date	Summary of Changes
TRM 0803.1	02/27/2014	• Reformatted; revised for inclusion of MPEG-DASH

4 Acronyms

AAC Advanced Audio Coding
FTP File Transfer Protocol
HTTP Hypertext Transfer Protocol
IBMF ISO Base Media Format
IP Internet Protocol
RTP Real Time Protocol

RTCP	Real Time Control Protocol
RTSP	Real Time Streaming Protocol
SAP	Session Announcement Protocol
SDP	Session Description Protocol
TCP	Transmission Control Protocol
TS	MPEG-2 Transport Stream
UDP	User Datagram Protocol
URL	Uniform Resource Locator

5 Introduction

The motion imagery produced at the Command and Control (C2) station of a motion imagery collection system (collection platform and control station) is considered the exploitation quality motion imagery for the system. The typical exploitation quality motion imagery is generated at MISP MISM level 3 [1] and greater. The motion imagery produced at these levels can exceed the bandwidth of the network delivering motion imagery to users at the edges. To meet network constraints motion imagery asset tradeoffs are required; either the imagery must be reduced in spatial resolution, decreased in temporal rate, or reduced in fidelity. In some cases a combination of these reductions is necessary. Alternative measures to meet bandwidth constraints include reducing the quantity of metadata, and transcoding the imagery using different compression. The motion imagery quality of a reduced-bandwidth motion imagery asset is typically in the range of MISP MISM levels 0 to 2 [1]. Because salient features in the original imagery may be lost in such highly compressed imagery, it should not be relied on for exploitation and targeting purposes.

Characterizing a channel as “low bandwidth” usually suggests a narrow channel or pipe, where information flow will not support real-time motion imagery, such as motion imagery. But low bandwidth also can include higher-capacity channels where interference may be high causing a reduced data throughput, or excessive network traffic incurs collisions and lost data. What is relevant is the overall data throughput rather than a high versus low bandwidth designation. Knowledge of network conditions is thus important in determining what technologies are best choices to achieve a desired level of performance.

This TRM is an aide for choosing technologies that meet design criteria for a given level of user performance within the MISB portfolio of adopted standards and practices.

6 Low Bandwidth Motion Imagery Components

Design objectives and criteria will suggest some particular combination of media types, delivery protocol, and distribution performance that will shape the user experience (see Figure 1). The network type (i.e. IP, RF link, etc.), the level of interactivity with the content (i.e. TiVo functions, real-time, etc.), and a means for discovery of content (pull versus push) are all factors in choosing an optimal delivery protocol for an application. MPEG-2 Transport Stream, RTP, and HTTP/FTP are delivery protocols for the motion imagery, metadata, and other media.

MISB-compliant motion imagery is either MPEG-2 or H.264 compressed and associated with KLV metadata structured according to MISB ST 1402 [2]. Audio and other annotation data may be included, but is not specified here.

6.1 Motion Imagery Encoding

The approved motion imagery encoding formats for low-bandwidth motion imagery is specified in the MISP RECOMMENDED PRACTICE 9720e–MISM, Low Bandwidth Motion Imagery Levels L1.2-L1.0. The MISB approved compression for these MISM levels is H.264/AVC, which supports data rates between 384 and 56 Kbits/sec.

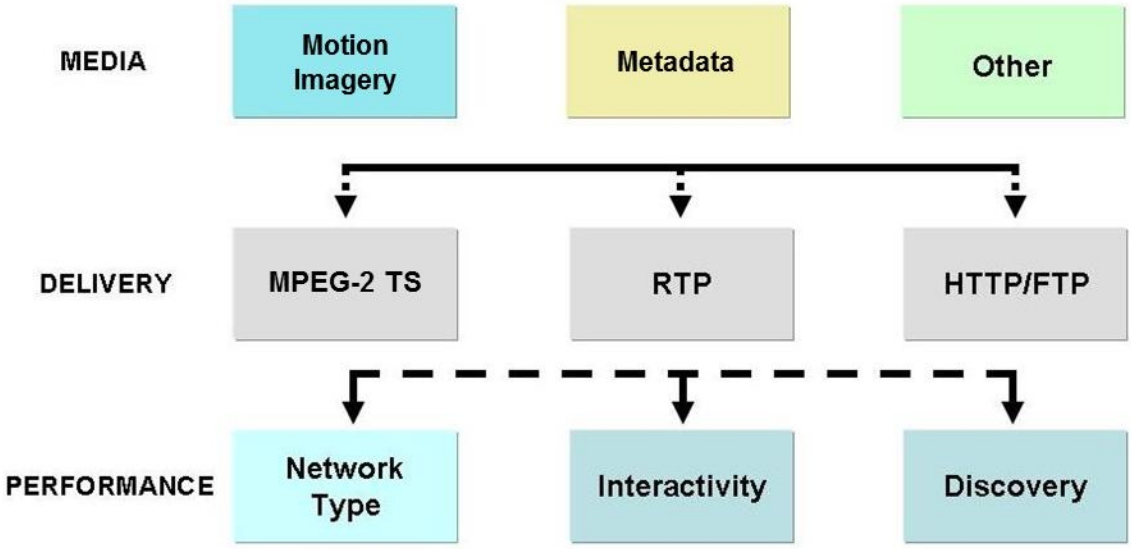

Figure 1: Elements That Shape a User Experience

6.2 Motion imagery Distribution Protocols

The approved distribution protocols for real-time delivery of low-bandwidth motion imagery are MPEG-2 Transport Stream (TS) and Real Time Protocol (RTP). The choice in delivery protocol is largely a matter of network type, desired interactivity, performance, and to a lesser degree the discovery means for available content.

6.2.1 MPEG-2 Transport Stream

Together, the MPEG-2 Transport Stream protocol [3] and the MISB define the rules for transporting motion imagery, audio, metadata, and other supporting data within the stream. MPEG-2 TS carries all media components (motion imagery, metadata, other) that are associated with a collect within one unified transport package.

6.2.1.1 Network Type

MPEG-2 TS is designed for constant delay networks and robustness in the presence of channel noise and interference, such as is typical in broadcast application. Thus, for transmission mediums that are point-to-point circuit-based, such as an RF link, MPEG-2 TS is a preferred protocol. MPEG-2 TS is frequently used to carry multimedia over *packet-based* networks, where best performance is observed when these networks are dedicated with little conflicting traffic and few network hops. MPEG-2 TS does not, however, perform well over a high-traffic, shared channel such as the public internet where data may travel through multiple network hops comprised of possibly disparate networks.

For channels that suffer lost data from interference Forward Error Correction (FEC) techniques may further ensure adequate performance at the expense of additional overhead. Suitable FEC methods include only those that are transparent (systematic) to receivers not designed to decode the additional protection codes.

6.2.1.2 Interactivity

The delivery of MPEG-2 transport stream using UDP/IP over circuit-based or packet-based networks does not afford viewer control to pause, rewind, or skip through the stream, unless the stream is further buffered or saved at the client.

6.2.1.3 Discovery

The discovery of content transported over MPEG-2 transport stream requires knowledge of the source; either a frequency to tune if broadcast, or the IP address of the sender if sent via packet-based means.

6.2.2 RTP/RTCP

Real Time Protocol (RTP) is a real-time end-to-end transport protocol developed for packet-based networks [4]. While RTP is the carrier of the data, RTCP (Real Time Control Protocol) manages the session by providing feedback on the quality of the data delivered and information about session participants. RTP rides on top of UDP/IP and a RTP session is composed of a RTP port number (UDP port), a RTCP port number (consecutive UDP port) and the participant's IP address. Use of RTP for motion imagery and metadata can be found in MISB ST 0804 [5].

6.2.2.1 Network Type

RTP is designed for packet-based networks where delivery cannot be guaranteed because of packet loss, jitter, and latency. Responsibility for data delivery is at the application level. RTP thus requires smart application endpoints to manage the data for optimal delivery. RTP is a preferred protocol in high-traffic shared networks, but is also applicable to any packet-based network.

Unlike MPEG-2 TS that carries all the media components associated with a collect within one package, only one media component is sent per RTP/RTCP connection. For example, motion imagery would require one RTP and one RTCP connection, while metadata would require its own separate RTP/RTCP connection. Synchronization of the various streams is re-established at the client through timestamps included in each stream, and a clock reference and timestamp

supplied in the respective RTCP channels. Xon2 may be carried over RTP; although this incurs additional bandwidth and is less seldom done.

6.2.2.2 Interactivity

Real Time Streaming Protocol (RTSP) [6] provides for setup, tear-down, and interactive user control of a RTP session. Functions such as pause, rewind, skip, stop, and play are supported by moving through timestamps within the content. This interactivity requires that a streaming server be employed to serve the content. A streaming server is software that formats the content to support interactivity by providing additional information such as hinting tracks. It also helps guard against perturbations incurred in transmission by shuffling the data to non-adjacent packets. Content is organized to support lower resolution and fidelity streams under changing network conditions, like sudden drops in channel bandwidth.

6.2.2.3 Discovery

RTP and RTCP connections are announced (made known) through the guidance of yet other protocols. One commonly used protocol is the Session Announcement Protocol (SAP) which further relies on the Session Description Protocol (SDP) [7]. SAP periodically issues (announces) IP addresses that clients can listen for and respond to. SDP provides the description of session information for the receivers to connect.

6.2.3 HTTP

If real-time delivery is not required the preferred method of delivery is file transfer over HTTP/TCP/IP. HTTP (Hypertext Transfer Protocol) [8] is a network protocol built upon TCP used to connect a web interface to a server. The delivery of data is thus guaranteed by virtue of TCP reliance, and assured transversal through network fire-walls.

6.2.3.1 Network Type

A client requests data by sending both its IP address and port number to port 80 of an addressed server. Port 80 is the default "open" port available for HTTP traffic. This assures that data can get through existing firewalls that might otherwise block reception. The advantage of HTTP is that the data is guaranteed to be delivered, since HTTP uses TCP/IP as the underlying IP protocol. An advantage over RTP, HTTP relies on common web servers to source the motion imagery instead of more costly streaming servers.

6.2.3.2 Interactivity

Three types of HTTP file transfer are common: *adaptive bitrate streaming*, *progressive download* and *download*. Adaptive bitrate streaming comes in several flavors: HTTP Live Streaming HLS (Apple), HTTP Dynamic streaming HDS (Adobe), Sure Streaming (Microsoft), and Dynamic Streaming over HTTP (MPEG-DASH). All are similar in that they deliver short segments (2-10 seconds) of content encoded in a number of different bitrates over TCP/IP. The client makes requests for a particular bitrate version of the content based on available bandwidth. In this way, an uninterrupted flow of content is possible; image quality changes as a function of bandwidth availability. A manifest file is communicated from server to client that indicates

parameters of the streamed session, such as compression type and file names for the segments. Two containers support MPEG-DASH: MPEG-2 TS and the ISO Base Media Format (IBMF).

Progressive download provides for file viewing *before* the entire file is downloaded; usually after 10-15% of the file has been received. Media client players must support progressive download to take advantage of this. In *download*, viewing is not available until the entire file is received.

Dependent on the client the user can interact with the content analogous to TiVo. Content is persistent at the client until either a new stream overwrites it, or the player application is terminated. The user has the ability to continually review the content and store the content unlike that for real-time streams.

6.2.3.3 Discovery

Content is accessed through a client web browser by clicking on its URL (Uniform Resource Locator). Content is made known through publication of a listing of URL's on a web site.

6.3 Metadata

The approved metadata constructs are encoded in KLV according to SMPTE ST 336 [9]. Metadata may be carried either asynchronously or synchronously in MPEG-2 Transport Stream (see MISB ST 1402). Metadata may also be carried in RTP as specified in RFC 6597 [[10]. The IBMF container used in MPEG-DASH supports the carriage of metadata, although there is no direct support for KLV-encoded metadata.

6.4 Audio

Audio may be carried with an MPEG-2 Transport Stream, or as RTP. Allowed formats for audio can be found in MISB ST 1001 [11].

6.5 Summary

Motion Imagery for Low Bandwidth Guide		
Motion Imagery	H264/AVC	Maps to MPEG-2 TS and RTP
Profile	MISM L1.2-L1.0	MISP
Metadata	KLV	Mappings to CoT
Audio	MPEG-1/2 Layer II MPEG-2 AAC-LC	Options: 2 channels, SBR
Transport Protocol	MPEG-2 Transport Stream	RF, dedicated, or low traffic packet-based networks
	RTP/RTCP	Packet-based
	MPEG-2 TS/IBMF	HTTP adaptive bitrate streaming